

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Lin, Lih-Ling
Graham, James
- (ii) TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
BINDING
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Brown, Scott A.
 - (B) REGISTRATION NUMBER: 32,724
 - (C) REFERENCE/DOCKET NUMBER: GI5258
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (617) 498-8224
 - (B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1571 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 2..529
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

G ATC CCC AGG GTG GAC CTC CGG GTG TGG CAG GAC TGC TGT GAA GAC
 Ile Pro Arg Val Asp Leu Arg Val Trp Gln Asp Cys Cys Glu Asp
 1 5 10 15

46

TGT	AGG	ACC	AGG	GGG	CAG	TTC	AAT	GCC	TTT	TCC	TAT	CAT	TTC	CGA	GGC		94
Cys	Arg	Thr	Arg	Gly 20	Gln	Phe	Asn	Ala	Phe 25	Ser	Tyr	His	Phe	Arg 30	Gly		
AGA	CGG	TCT	CTT	GAG	TTC	AGC	TAC	CAG	GAG	GAC	AAG	CCG	ACC	AAG	AAA		142
Arg	Arg	Ser	Leu	Glu	Phe	Ser	Tyr	Gln 40	Glu	Asp	Lys	Pro	Thr 45	Lys	Lys		
ACA	AGA	CCA	CGG	AAA	ATA	CCC	AGT	GTT	GGG	AGA	CAG	GGG	GAA	CAT	CTC		190
Thr	Arg	Pro	Arg	Lys	Ile	Pro	Ser	Val	Gly	Arg	Gln	Gly	Glu	His	Leu		
		50					55					60					
AGC	AAC	AGC	ACC	TCA	GCC	TTC	AGC	ACA	CGC	TCA	GAT	GCA	TCT	GGG	ACA		238
Ser	Asn	Ser	Thr	Ser	Ala	Phe	Ser	Thr	Arg	Ser	Asp	Ala	Ser	Gly	Thr		
	65					70					75						
AAT	GAC	TTC	AGA	GAG	TTT	GTT	CTG	GAA	ATG	CAG	AAG	ACC	ATC	ACA	GAC		286
Asn	Asp	Phe	Arg	Glu	Phe	Val	Leu	Glu	Met	Gln	Lys	Thr	Ile	Thr	Asp		
	80				85					90					95		
CTC	AGA	ACA	CAG	ATA	AAG	AAA	CTT	GAA	TCA	CGG	CTC	AGT	ACC	ACA	GAG		334
Leu	Arg	Thr	Gln	Ile	Lys	Lys	Leu	Glu	Ser	Arg	Leu	Ser	Thr	Thr	Glu		
				100					105						110		
TGC	GTG	GAT	GCC	GGG	GGC	GAA	TCT	CAC	GCC	AAC	AAC	ACC	AAG	TGG	AAA		382
Cys	Val	Asp	Ala	Gly	Gly	Glu	Ser	His	Ala	Asn	Asn	Thr	Lys	Trp	Lys		
			115					120					125				
AAA	GAT	GCA	TGC	ACC	ATT	TGT	GAA	TGC	AAA	GAC	GGG	CAG	GTC	ACC	TGC		430
Lys	Asp	Ala	Cys	Thr	Ile	Cys	Glu	Cys	Lys	Asp	Gly	Gln	Val	Thr	Cys		
		130					135					140					
TTC	GTG	GAA	GCT	TGC	CCC	CCT	GCC	ACC	TGT	GCT	GTC	CCC	GTG	AAC	ATC		478
Phe	Val	Glu	Ala	Cys	Pro	Pro	Ala	Thr	Cys	Ala	Val	Pro	Val	Asn	Ile		
		145				150					155						
CCA	GGG	GCC	TGC	TGT	CCA	GTC	TGC	TTA	CAG	AAG	AGG	GCG	GAG	GAA	AAG		526
Pro	Gly	Ala	Cys	Cys	Pro	Val	Cys	Leu	Gln	Lys	Arg	Ala	Glu	Glu	Lys		
					165					170					175		
CCC	TAGGCTCCTG	GGAGGCTCCT	CAGAGTTTGT	CTGCTGTGCC	ATCGTGAGAT												579
Pro																	
CGGGTGGCCG	ATGGCAGGGA	GCTGCGGACT	GCAGACCAGG	AAACACCCAG	AACTCGTGAC												639
ATTCATGAC	AACGTCCAGC	TGGTGCTGTT	ACAGAAGGCA	GTGCAGGAGG	CTTCCAACCA												699
GAGCATCTGC	GGAGAAGGAG	GCACAGCAGG	TGCCTGAAGG	GAAGCAGGCA	GGAGTCCTAG												759
CTTCACGTTA	GACTTCTCAG	GTTTTTATTT	AATTCTTTTA	AAATGAAAAA	TTGGTGCTAC												819
TATTAAATTG	CACAGTTGAA	TCATTTAGGC	GCCTAAATTG	ATTTTGCCTC	CCAACACCAT												879
TTCTTTTTAA	ATAAAGCAGG	ATACCTCTAT	ATGTCAGCCT	TGCCTTGTTT	AGATGCCAGG												939
AGCCGGCAGA	CCTGTCACCC	GCAGGTGGGG	TGAGTCTCGG	AGCTGCCAGA	GGGGCTCACC												999
GAAATCGGGG	TTCCATCAC	AGCTATGTTT	AAAAAGAAAA	TTGGTGTTTG	CCAAACGGAA												1059
CAGAACCTTT	GATGAGAGCG	TTCACAGGGA	CACTGTCTGG	GGGTGCAGTG	CAAGCCCCCG												1119
GCCTCTTCCC	TGGGAACCTC	TGAACTCCTC	CTTCCTCTGG	GCTCTCTGTA	ACATTTCACC												1179
ACACGTCAGC	ATCTAATCCC	AAGACAAACA	TTCCCGCTGC	TCGAAGCAGC	TGTATAGCCT												1239
GTGACTCTCC	GTGTGTCAGC	TCCTTCCACA	CCTGATTAG	ACATTTCATAA	GCCACATTTA												1299

GAAACAGGTT TGCTTTTCAGC TGTCACCTGC ACACATACTG CCTAGTTGTG AACCAAATGT 1359
 GAAAAAACCT CCTTCATCCC ATTGTGTATC TGATACCTGC CGAGGGCCAA GGGTGTGTGT 1419
 TGACAACGCC GCTCCCAGCC GGCCCTGGTT GCGTCCACGT CCTGAACAAG AGCCGCTTCC 1479
 GGATGGCTCT TCCCAAGGGA GGAGGAGCTC AAGTGTCTGGG AACTGTCTAA CTTCAGGTTG 1539
 TGTGAGTGCG TTAAAAAAA AAAAAAAA AA 1571

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ile Pro Arg Val Asp Leu Arg Val Trp Gln Asp Cys Cys Glu Asp Cys
 1 5 10 15
 Arg Thr Arg Gly Gln Phe Asn Ala Phe Ser Tyr His Phe Arg Gly Arg
 20 25 30
 Arg Ser Leu Glu Phe Ser Tyr Gln Glu Asp Lys Pro Thr Lys Lys Thr
 35 40 45
 Arg Pro Arg Lys Ile Pro Ser Val Gly Arg Gln Gly Glu His Leu Ser
 50 55 60
 Asn Ser Thr Ser Ala Phe Ser Thr Arg Ser Asp Ala Ser Gly Thr Asn
 65 70 75 80
 Asp Phe Arg Glu Phe Val Leu Glu Met Gln Lys Thr Ile Thr Asp Leu
 85 90 95
 Arg Thr Gln Ile Lys Lys Leu Glu Ser Arg Leu Ser Thr Thr Glu Cys
 100 105 110
 Val Asp Ala Gly Gly Glu Ser His Ala Asn Asn Thr Lys Trp Lys Lys
 115 120 125
 Asp Ala Cys Thr Ile Cys Glu Cys Lys Asp Gly Gln Val Thr Cys Phe
 130 135 140
 Val Glu Ala Cys Pro Pro Ala Thr Cys Ala Val Pro Val Asn Ile Pro
 145 150 155 160
 Gly Ala Cys Cys Pro Val Cys Leu Gln Lys Arg Ala Glu Glu Lys Pro
 165 170 175

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1088 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..961

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

G AAA AAA GGA GGT AAA ACA GAA CAG GAT GGC TAT CAG AAA CCC ACC	46
Lys Lys Gly Gly Lys Thr Glu Gln Asp Gly Tyr Gln Lys Pro Thr	
1 5 10 15	
AAC AAA CAC TTC ACG CAG AGT CCC AAG AAG TCA GTG GCC GAC CTG CTG	94
Asn Lys His Phe Thr Gln Ser Pro Lys Lys Ser Val Ala Asp Leu Leu	
20 25 30	
GGG TCC TTT GAA GGC AAA CGA AGA CTC CTT CTG ATC ACT GCT CCC AAG	142
Gly Ser Phe Glu Gly Lys Arg Arg Leu Leu Leu Ile Thr Ala Pro Lys	
35 40 45	
GCT GAG AAC AAT ATG TAT GTG CAA CAA CGT GAT GAA TAT CTG GAA AGT	190
Ala Glu Asn Asn Met Tyr Val Gln Gln Arg Asp Glu Tyr Leu Glu Ser	
50 55 60	
TTC TGC AAG ATG GCT ACC AGG AAA ATC TCT GTG ATC ACC ATC TTC GGC	238
Phe Cys Lys Met Ala Thr Arg Lys Ile Ser Val Ile Thr Ile Phe Gly	
65 70 75	
CCT GTC AAC AAC AGC ACC ATG AAA ATC GAC CAC TTT CAG CTA GAT AAT	286
Pro Val Asn Asn Ser Thr Met Lys Ile Asp His Phe Gln Leu Asp Asn	
80 85 90 95	
GAG AAG CCC ATG CGA GTG GTG GAT GAT GAA GAC TTG GTA GAC CAG CGT	334
Glu Lys Pro Met Arg Val Val Asp Asp Glu Asp Leu Val Asp Gln Arg	
100 105 110	
CTC ATC AGC GAG CTG AGG AAA GAG TAC GGA ATG ACC TAC AAT GAC TTC	382
Leu Ile Ser Glu Leu Arg Lys Glu Tyr Gly Met Thr Tyr Asn Asp Phe	
115 120 125	
TTC ATG GTG CTA ACA GAT GTG GAT CTG AGA GTC AAG CAA TAC TAT GAG	430
Phe Met Val Leu Thr Asp Val Asp Leu Arg Val Lys Gln Tyr Tyr Glu	
130 135 140	
GTA CCA ATA ACA ATG AAG TCT GTG TTT GAT CTG ATC GAT ACT TTC CAG	478
Val Pro Ile Thr Met Lys Ser Val Phe Asp Leu Ile Asp Thr Phe Gln	
145 150 155	
TCC CGA ATC AAA GAT ATG GAG AAG CAG AAG AAG GAG GGC ATT GTT TGC	526
Ser Arg Ile Lys Asp Met Glu Lys Gln Lys Lys Glu Gly Ile Val Cys	
160 165 170 175	
AAA GAG GAA GTT GGG GGA GTG TTA GAA CTG TTC CCA ATT AAT GGG AGC	574
Lys Glu Glu Val Gly Gly Val Leu Glu Leu Phe Pro Ile Asn Gly Ser	
180 185 190	
TCT GTT GTT GAG CGA GAA GAC GTA CCA GCC CAT TTG GTG AAA GAC ATT	622
Ser Val Val Glu Arg Glu Asp Val Pro Ala His Leu Val Lys Asp Ile	
195 200 205	
CGT AAC TAT TTT CAA GTG AGC CCG GAG TAC TTC TCC ATG CTT CTA GTC	670
Arg Asn Tyr Phe Gln Val Ser Pro Glu Tyr Phe Ser Met Leu Leu Val	
210 215 220	
GGA AAA GAC GGA AAT GTC AAA TCC TGG TAT CCT TCC CCA ATG TGG TCC	718
Gly Lys Asp Gly Asn Val Lys Ser Trp Tyr Pro Ser Pro Met Trp Ser	
225 230 235	

ATG GTG ATT GTG TAC GAT TTA ATT GAT TCG ATG CAA CTT CGG AGA CAG	766
Met Val Ile Val Tyr Asp Leu Ile Asp Ser Met Gln Leu Arg Arg Gln	
240 245 250 255	
GAA ATG GCG ATT CAG CAG TCA CTG GGG ATG CGC TGC CAG AAG ATG AGT	814
Glu Met Ala Ile Gln Gln Ser Leu Gly Met Arg Cys Gln Lys Met Ser	
260 265 270	
ATG CAG GCT ATG GTT ACC ATA GTT ACC ACC AAG GAT ACC AGG ATG GTT	862
Met Gln Ala Met Val Thr Ile Val Thr Thr Lys Asp Thr Arg Met Val	
275 280 285	
ACC AGG ATG ACT ACC GTC ATC ATG AGA GTT ATC ACC ATG GAT ACC CTT	910
Thr Arg Met Thr Thr Val Ile Met Arg Val Ile Thr Met Asp Thr Leu	
290 295 300	
ACT GAG CAG AAA TAT GTA ACC TTA GAC TCA GCC AGT TTC CTC TGC AGC	958
Thr Glu Gln Lys Tyr Val Thr Leu Asp Ser Ala Ser Phe Leu Cys Ser	
305 310 315	
TGC TAAAACTACA TGTGGCCAGC TCCATTCTTC CACACTGCGT ACTACATTTT	1011
Cys	
320	
CTGCCTTTTT CTTCAGTGT TTTTCTAAGA CTAAATAAAT AGCAAACCTTT CACCTAAAAA	1071
AAAAAAAAAA AAAAAA	1088

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 320 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Lys Lys Gly Gly Lys Thr Glu Gln Asp Gly Tyr Gln Lys Pro Thr Asn
1      5      10      15
Lys His Phe Thr Gln Ser Pro Lys Lys Ser Val Ala Asp Leu Leu Gly
20      25      30
Ser Phe Glu Gly Lys Arg Arg Leu Leu Ile Thr Ala Pro Lys Ala
35      40      45
Glu Asn Asn Met Tyr Val Gln Gln Arg Asp Glu Tyr Leu Glu Ser Phe
50      55      60
Cys Lys Met Ala Thr Arg Lys Ile Ser Val Ile Thr Ile Phe Gly Pro
65      70      75      80
Val Asn Asn Ser Thr Met Lys Ile Asp His Phe Gln Leu Asp Asn Glu
85      90      95
Lys Pro Met Arg Val Val Asp Asp Glu Asp Leu Val Asp Gln Arg Leu
100     105     110
Ile Ser Glu Leu Arg Lys Glu Tyr Gly Met Thr Tyr Asn Asp Phe Phe
115     120     125
Met Val Leu Thr Asp Val Asp Leu Arg Val Lys Gln Tyr Tyr Glu Val
130     135     140

```

Pro Ile Thr Met Lys Ser Val Phe Asp Leu Ile Asp Thr Phe Gln Ser
 145 150 155 160
 Arg Ile Lys Asp Met Glu Lys Gln Lys Lys Glu Gly Ile Val Cys Lys
 165 170 175
 Glu Glu Val Gly Gly Val Leu Glu Leu Phe Pro Ile Asn Gly Ser Ser
 180 185 190
 Val Val Glu Arg Glu Asp Val Pro Ala His Leu Val Lys Asp Ile Arg
 195 200 205
 Asn Tyr Phe Gln Val Ser Pro Glu Tyr Phe Ser Met Leu Leu Val Gly
 210 215 220
 Lys Asp Gly Asn Val Lys Ser Trp Tyr Pro Ser Pro Met Trp Ser Met
 225 230 235 240
 Val Ile Val Tyr Asp Leu Ile Asp Ser Met Gln Leu Arg Arg Gln Glu
 245 250 255
 Met Ala Ile Gln Gln Ser Leu Gly Met Arg Cys Gln Lys Met Ser Met
 260 265 270
 Gln Ala Met Val Thr Ile Val Thr Thr Lys Asp Thr Arg Met Val Thr
 275 280 285
 Arg Met Thr Thr Val Ile Met Arg Val Ile Thr Met Asp Thr Leu Thr
 290 295 300
 Glu Gln Lys Tyr Val Thr Leu Asp Ser Ala Ser Phe Leu Cys Ser Cys
 305 310 315 320

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1759 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 2..754

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

C AAA AAC TTC TTC CTG ACG AAT CGC GCC AGG GAG CGC TCA GAC ACC	46
Lys Asn Phe Phe Leu Thr Asn Arg Ala Arg Glu Arg Ser Asp Thr	
1 5 10 15	
TTC ATC AAC CTC CGG GAG GTG CTC AAC CGC TTC AAG CTG CCG CCA GGA	94
Phe Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly	
20 25 30	
GAG TAC ATT CTC GTG CCT TCC ACC TTC GAA CCC AAC AAG GAT GGG GAT	142
Glu Tyr Ile Leu Val Pro Ser Thr Phe Glu Pro Asn Lys Asp Gly Asp	
35 40 45	

TTC Phe	TGC Cys	ATC Ile	CGG Arg	GTC Val	TTT Phe	TCT Ser	GAA Glu	AAA Lys	AAA Lys	GCT Ala	GAC Asp	TAC Tyr	CAA Gln	GCT Ala	ATC Val	190
50 55 60																
GAT Asp	GAT Asp	GAA Glu	ATC Ile	GAG Glu	GCC Ala	AAT Asn	CTT Leu	GAA Glu	GAG Glu	TTC Phe	GAC Asp	ATC Ile	AGC Ser	GAG Glu	GAT Asp	238
65 70 75																
GAC Asp	ATT Ile	GAT Asp	GAT Asp	GGA Gly	TTC Phe	AGG Arg	AGA Arg	CTG Leu	TTT Phe	GCC Ala	CAG Gln	TTG Leu	GCA Ala	GGA Gly	GAG Glu	286
80 85 90 95																
GAT Asp	GCG Ala	GAG Glu	ATC Ile	TCT Ser	GCC Ala	TTT Phe	GAG Glu	CTG Leu	CAG Gln	ACC Thr	ATC Ile	CTG Leu	AGA Arg	AGG Arg	GTT Val	334
100 105 110																
CTA Leu	GCA Ala	AAG Lys	CGC Arg	CAA Gln	GAT Asp	ATC Ile	AAG Lys	TCA Ser	GAT Asp	GGC Gly	TTC Phe	AGC Ser	ATC Ile	GAG Glu	ACA Thr	382
115 120 125																
TGC Cys	AAA Lys	ATT Ile	ATG Met	GTT Val	GAC Asp	ATG Met	CTA Leu	GAT Asp	TCG Ser	GAC Asp	GGG Gly	AGT Ser	GGC Gly	AAG Lys	CTG Leu	430
130 135 140																
GGG Gly	CTG Leu	AAG Lys	GAG Glu	TTC Phe	TAC Tyr	ATT Ile	CTC Leu	TGG Trp	ACG Thr	AAG Lys	ATT Ile	CAA Gln	AAA Lys	TAC Tyr	CAA Gln	478
145 150 155																
AAA Lys	ATT Ile	TAC Tyr	CGA Arg	GAA Glu	ATC Ile	GAC Asp	GTT Val	GAC Asp	AGG Arg	TCT Ser	GGT Gly	ACC Thr	ATG Met	AAT Asn	TCC Ser	526
160 165 170 175																
TAT Tyr	GAA Glu	ATG Met	CGG Arg	AAG Lys	GCA Ala	TTA Leu	GAA Glu	GAA Glu	GCA Ala	GGT Gly	TTC Phe	AAG Lys	ATG Met	CCC Pro	TGT Cys	574
180 185 190																
CAA Gln	CTC Leu	CAC His	CAA Gln	GTC Val	ATC Ile	GTT Val	GCT Ala	CGG Arg	TTT Phe	GCA Ala	GAT Asp	GAC Asp	CAG Gln	CTC Leu	ATC Ile	622
195 200 205																
ATC Ile	GAT Asp	TTT Phe	GAT Asp	AAT Asn	TTT Phe	GTT Val	CGG Arg	TGT Cys	TTG Leu	GTT Val	CGG Arg	CTG Leu	GAA Glu	ACG Thr	CTA Leu	670
210 215 220																
TTC Phe	AAG Lys	ATA Ile	TTT Phe	AAG Lys	CAG Gln	CTG Leu	GAT Asp	CCC Pro	GAG Glu	AAT Asn	ACT Thr	GGA Gly	ACA Thr	ATA Ile	GAG Glu	718
225 230 235																
CTC Leu	GAC Asp	CTT Leu	ATC Ile	TCT Ser	TGG Trp	CTC Leu	TGT Cys	TTC Phe	TCA Ser	GTA Val	CTT Leu	TGAAGTTATA				764
240 245 250																
ACTAATCTGC CTGAAGACTT CTCATGATGG AAAATCAGCC AAGGACTAAG CTTCCATAGA 824																
AATACACTTT GTATCTGGAC CTCAAAATTA TGGAACATT TACTTAAACG GATGATCATA 884																
GCTGAAAATA ATGATACTGT CAATTTGAGA TAGCAGAAGT TTCACACATC AAAGTAAAG 944																
ATTTGCATAT CATTATACTA AATGCAAATG AGTCGCTTAA CCCTTGACAA GGTCAAAGAA 1004																
AGCTTTAAAT CTGTAAATAG TATACACTTT TTACTTTTAC ACACTTTCCT GTTCATAGCA 1064																
ATATTAAATC AGGAAAAAAA AATGCAGGGA GGTATTTAAC AGCTGAGCAA AAACATTGAG 1124																
TCGCTCTCAA AGGACACGAG GCCCTTGGCA GGAATATTT AAAGCAACTT CAGTTTAAA 1184																
ATGCAGCTGT TGATTCTACC AAACAACAGT CCAAGATTAC CATTTCCTAT GAGCCAACTG 1244																

GGAAACATGG TATATCATGA AGTAATCTTG TCAAGGCATC TGGAGAGTCC AGGAGAGAAG 1304
 ACTCACCTCT GTCGCTTGGG TTAAACAAGA GACAGGTTTT GTAGAATATT GATTGGTAAT 1364
 AGTAAATCGT TCTCCTTACA ATCAAGTTCT TGACCCTATT CGGCCTTATA CATCTGGTCT 1424
 TACAAAGACC AAAGGGATCC TGCCTTGAT CAACTGAACC AGTATGCCAA AACCAGGCAT 1484
 CCAATTTGTA AACCAATTAT GATAAAGGAC AAAATAAGCT GTTTGCCACC TCAAACTTT 1544
 ATGAACTTCA CCACCACTAG TGTCTGTCCA TGGAGTTAGA GGGGACATCA CTTAGAAGTT 1604
 CTTATAGAAA GGACACAAGT TTGTTTCCTG GCTTTACCTT GGGAAAATGC TAGCAACATT 1664
 ATAGAAATTT TGCCTTGTTG CCTTATCTTC TTCCAAATGT ACTGTAAAT AAAAATAAAG 1724
 GGTTACCCCA TGCAATCAAA AAAAAAAAAA AAAAA 1759

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Asn Phe Phe Leu Thr Asn Arg Ala Arg Glu Arg Ser Asp Thr Phe
 1 5 10 15
 Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly Glu
 20 25 30
 Tyr Ile Leu Val Pro Ser Thr Phe Glu Pro Asn Lys Asp Gly Asp Phe
 35 40 45
 Cys Ile Arg Val Phe Ser Glu Lys Lys Ala Asp Tyr Gln Ala Val Asp
 50 55 60
 Asp Glu Ile Glu Ala Asn Leu Glu Glu Phe Asp Ile Ser Glu Asp Asp
 65 70 75 80
 Ile Asp Asp Gly Phe Arg Arg Leu Phe Ala Gln Leu Ala Gly Glu Asp
 85 90 95
 Ala Glu Ile Ser Ala Phe Glu Leu Gln Thr Ile Leu Arg Arg Val Leu
 100 105 110
 Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr Cys
 115 120 125
 Lys Ile Met Val Asp Met Leu Asp Ser Asp Gly Ser Gly Lys Leu Gly
 130 135 140
 Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln Lys
 145 150 155 160
 Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser Tyr
 165 170 175
 Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Met Pro Cys Gln
 180 185 190

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 700 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

Met 1	Ala	Gly	Ile	Ala 5	Ala	Lys	Leu	Ala	Lys 10	Asp	Arg	Glu	Ala	Ala 15	Glu
Gly	Leu	Gly	Ser 20	His	Glu	Arg	Ala	Ile 25	Lys	Tyr	Leu	Asn	Gln 30	Asp	Tyr
Glu	Ala	Leu 35	Arg	Asn	Glu	Cys	Leu 40	Glu	Ala	Gly	Thr	Leu 45	Phe	Gln	Asp
Pro 50	Ser	Phe	Pro	Ala	Ile	Pro 55	Ser	Ala	Leu	Gly	Phe 60	Lys	Glu	Leu	Gly
Pro 65	Tyr	Ser	Ser	Lys 70	Thr	Arg	Gly	Met	Arg 75	Trp	Lys	Arg	Pro	Thr	Glu 80
Ile	Cys	Ala	Asp 85	Pro	Gln	Phe	Ile	Ile	Gly 90	Gly	Ala	Thr	Arg	Thr 95	Asp
Ile	Cys	Gln 100	Gly	Ala	Leu	Gly	Asp	Cys 105	Trp	Leu	Leu	Ala	Ala 110	Ile	Ala
Ser 115	Leu	Thr	Leu	Asn	Glu	Glu	Ile 120	Leu	Ala	Arg	Val 125	Val	Pro	Leu	Asn
Gln 130	Ser	Phe	Gln	Glu	Asn	Tyr 135	Ala	Gly	Ile	Phe	His 140	Phe	Gln	Phe	Trp
Gln 145	Tyr	Gly	Glu	Trp 150	Val	Glu	Val	Val	Val	Asp 155	Asp	Arg	Leu	Pro	Thr 160
Lys	Asp	Gly	Glu	Leu 165	Leu	Phe	Val	His	Ser 170	Ala	Glu	Gly	Ser	Glu 175	Phe
Trp	Ser	Ala	Leu 180	Leu	Glu	Lys	Ala	Tyr 185	Ala	Lys	Ile	Asn	Gly 190	Cys	Tyr
Glu	Ala 195	Leu	Ser	Gly	Gly	Ala	Thr 200	Thr	Glu	Gly	Phe	Glu 205	Asp	Phe	Thr

Gly	Gly	Ile	Ala	Glu	Trp	Tyr	Glu	Leu	Lys	Lys	Pro	Pro	Pro	Asn	Leu
	210					215					220				
Phe	Lys	Ile	Ile	Gln	Lys	Ala	Leu	Gln	Lys	Gly	Ser	Leu	Leu	Gly	Cys
225					230					235					240
Ser	Ile	Asp	Ile	Thr	Ser	Ala	Ala	Asp	Ser	Glu	Ala	Ile	Thr	Phe	Gln
				245					250					255	
Lys	Leu	Val	Lys	Gly	His	Ala	Tyr	Ser	Val	Thr	Gly	Ala	Glu	Glu	Val
			260					265					270		
Glu	Ser	Asn	Gly	Ser	Leu	Gln	Lys	Leu	Ile	Arg	Ile	Arg	Asn	Pro	Trp
		275					280					285			
Gly	Glu	Val	Glu	Trp	Thr	Gly	Arg	Trp	Asn	Asp	Asn	Cys	Pro	Ser	Trp
	290					295					300				
Asn	Thr	Ile	Asp	Pro	Glu	Glu	Arg	Glu	Arg	Leu	Thr	Arg	Arg	His	Glu
305					310					315					320
Asp	Gly	Glu	Phe	Trp	Met	Ser	Phe	Ser	Asp	Phe	Leu	Arg	His	Tyr	Ser
				325					330					335	
Arg	Leu	Glu	Ile	Cys	Asn	Leu	Thr	Pro	Asp	Thr	Leu	Thr	Ser	Asp	Thr
			340					345					350		
Tyr	Lys	Lys	Trp	Lys	Leu	Thr	Lys	Met	Asp	Gly	Asn	Trp	Arg	Arg	Gly
		355					360					365			
Ser	Thr	Ala	Gly	Gly	Cys	Arg	Asn	Tyr	Pro	Asn	Thr	Phe	Trp	Met	Asn
	370					375					380				
Pro	Gln	Tyr	Leu	Ile	Lys	Leu	Glu	Glu	Glu	Asp	Glu	Asp	Glu	Glu	Asp
385					390					395					400
Gly	Glu	Ser	Gly	Cys	Thr	Phe	Leu	Val	Gly	Leu	Ile	Gln	Lys	His	Arg
				405					410					415	
Arg	Arg	Gln	Arg	Lys	Met	Gly	Glu	Asp	Met	His	Thr	Ile	Gly	Phe	Gly
			420					425					430		
Ile	Tyr	Glu	Val	Pro	Glu	Glu	Leu	Ser	Gly	Gln	Thr	Asn	Ile	His	Leu
		435					440					445			
Ser	Lys	Asn	Phe	Phe	Leu	Thr	Asn	Arg	Ala	Arg	Glu	Arg	Ser	Asp	Thr
		450				455					460				
Phe	Ile	Asn	Leu	Arg	Glu	Val	Leu	Asn	Arg	Phe	Lys	Leu	Pro	Pro	Gly
465					470					475					480
Glu	Tyr	Ile	Leu	Val	Pro	Ser	Thr	Phe	Glu	Pro	Asn	Lys	Asp	Gly	Asp
				485					490					495	
Phe	Cys	Ile	Arg	Val	Phe	Ser	Glu	Lys	Lys	Ala	Asp	Tyr	Gln	Ala	Val
			500					505					510		
Asp	Asp	Glu	Ile	Glu	Ala	Asn	Leu	Glu	Glu	Phe	Asp	Ile	Ser	Glu	Asp
		515					520					525			
Asp	Ile	Asp	Asp	Gly	Val	Arg	Arg	Leu	Phe	Ala	Gln	Leu	Ala	Gly	Glu
	530					535					540				
Asp	Ala	Glu	Ile	Ser	Ala	Phe	Glu	Leu	Gln	Thr	Ile	Leu	Arg	Arg	Val
545					550					555					560

Leu Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr
 565 570 575
 Cys Lys Ile Met Val Asp Met Leu Asp Ser Asp Gly Ser Gly Lys Leu
 580 585 590
 Gly Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln
 595 600 605
 Lys Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser
 610 615 620
 Tyr Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Met Pro Cys
 625 630 635 640
 Gln Leu His Gln Val Ile Val Ala Arg Phe Ala Asp Asp Gln Leu Ile
 645 650 655
 Ile Asp Phe Asp Asn Phe Val Arg Cys Leu Val Arg Leu Glu Thr Leu
 660 665 670
 Phe Lys Ile Phe Lys Gln Leu Asp Pro Glu Asn Thr Gly Thr Ile Glu
 675 680 685
 Leu Asp Leu Ile Ser Trp Leu Cys Phe Ser Val Leu
 690 695 700